

RP SEQUENCE FROM N.A.
RC SPECIES=WHEAT; STRAIN=CV. MARDLER;
RX MEDLINE: 91329710.

DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE CYTOCHROME B6 (EC 1.10.99.1).
 GN PETB.
 OS ZEA MAYS (MAIZE).
 OC CHLOROPLAST.
 CC EUKARYOTA: PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONAE;
 CC CYPERALES; GRAMINEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 84210525.
 RA ROCK C.D., BARKAN A., TAYLOR W.C.;
 RL CURR. GENET. 12:69-77(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 95395841.
 RA MAIER R.M., NECKERMAN K., IGLIOI G.L., KOESSEL H.;
 RL J. MOL. BIOL. 251:614-628(1995).
 CC -1- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART
 CC OF THE CHLOROPLAST RESPIRATORY CHAIN.
 CC -1- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCTANTIN =
 CC PLASTOQUINONE + 2 REDUCED PLASTOCTANTIN.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN.
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,
 CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.
 CC -1- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.
 DR EMBL: X05422; G311718; -.
 DR EMBL: X86563; G902251; ALT_SEQ.
 DR PIR: S08592; CBZM6R.
 DR MAIZEDB: 56334; -.
 DR MENDEL: 4315; ZEAMA;PETB;1.
 DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
 KW ELECTRON TRANSPORT; HEME; CHLOROPLAST; PHOTOSYNTHESIS; TRANSMEMBRANE.
 FT METAL 86 86 IRON1 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 FT METAL 100 100 IRON2 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 FT METAL 187 187 IRON2 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 FT METAL 202 202 IRON1 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 SQ SEQUENCE 215 AA; 24180 MW; CBAB4737 CRC32;

Db 11 IFMSLSISMILTIY-LNNIFNPSMLLIYLISYSIYMSLMKFT-MCSMNSLLIMLILVF 68
| : : | | : : : : | : : : : | : : : : | : : : : |
Ov 147 IVASAGVMMI-TRYDLTHNERPAVILLMFSLVYKAFV-METEVHLCISGSHRIDARAV- 203

Db 69 LSMULM15
:::1:
Qy 204 VTGLLAL 210

RESULT 7
ID CYB6_SINP7 STANDARD; PRT; 215 AA.
AC Q54711;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CYTOCHROME B6 (EC 1.10.99.1).
GN PETB.
OS SYNECHOCOCCUS SP. (STRAIN PCC 7942) (ANACYSTIS NIDULANS R2).
OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
OC CYANOBACTERIA (BLUE-GREEN ALGAE); CHROCOCCALES.
RN [1]
RP SEQUENCE FROM N.A.
RA NEWMAN J., KARAKAYA H., SCANLAN D.J., CARR N.G., MANN N.R.;
RL SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART
CC OF THE CYANOBACTERIAL RESPIRATORY CHAIN.
CC -1- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCTANIN =
CC PLASTOQUINONE + 2 REDUCED PLASTOCTANIN.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,
CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.
CC -1- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.
DR ENBL: U33285; G988289; -.
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
KW ELECTRON TRANSPORT; HEME; PHOTOSYNTHESIS; TRANSMEMBRANE.
FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND)
FT (BY SIMILARITY).
FT METAL 100 100 IRON 2 (HEME AXIAL LIGAND)
FT (BY SIMILARITY).
FT METAL 187 187 IRON 2 (HEME AXIAL LIGAND)
FT (BY SIMILARITY).
FT METAL 202 202 IRON 1 (HEME AXIAL LIGAND)
FT (BY SIMILARITY).
SQ SEQUENCE 215 AA; 24357 MW; 38890BD7 CRC32;

Query Match 5.5%; Score 107; DB 1; Length 215;
Best Local Similarity 23.7%; Pred. No. 7.69e-02;
Matches 18; Conservative 20; Mismatches 32; Indels 6; Gaps 6;

Db 57 YIKPTVAEATSSVQIMNVNPGWLI-R-SI-HNWSAMVLMVLFVRLVLTGCFKRP 113
:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|
Qy 132 YIQMSIDNSILSVHTVSAQV-WNITRYDLYNFRPAVL-LMFLSVTKAFMETFVHL 189
:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|
Db 114 REL-TWTVGVNAVIT 128
:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|
Qy 190 CSLSGNSWRLDARAVVT 205
:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|

RESULT 8
ID CYB6_CHLRE STANDARD; PRT; 215 AA.
AC Q00471;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CYTOCHROME B6 (EC 1.10.99.1).
GN PETB.
OS CHLAMYDOMONAS REINHARDTII.
OC CHLOROPLAST.
OC EUKARYOTA; PLANTA; PHYCOPHYTA; CHLOROPHYTA (GREEN ALGAE);
OC CHLOROPHYCEAE; VOLVOCALES; CHLAMYDOMONADACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92256821.
RA HUANG C., LIU X.-Q.;
RL PLANT MOL. BIOL. 18:985-988(1992).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN-137C;
RX MEDLINE: 91285146.
RA BUESCHLEN S., CHOQUET Y., KURAS R., WOLLMAN F.A.;
RL FEBS LETT. 284:257-262(1991).
RN [3]
RP CHARACTERIZATION.
RC STRAIN-WT12;
RX MEDLINE: 96094329.
RA PIERRE Y., BREYTON C., KRAMER D., POPOUT J.-L.;
RL J. BIOL. CHEM. 270:29342-29349(1995).
CC -1- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART
CC OF THE CHLOROPLASTIC RESPIRATORY CHAIN.
CC -1- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCTANIN =
CC PLASTOQUINONE + 2 REDUCED PLASTOCTANIN.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,
CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.
CC -1- PM: THE N-TERMINAL IS BLOCKED.
CC -1- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.
DR ENBL: G62905; G12489; -.
DR ENBL: X72918; G288909; -.
DR PIR: S16917; S16917.
DR PIR: S21253; S21253.
DR MENDEL: 4312; CHLRE; PETB; 1.
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
KW ELECTRON-TRANSPORT; HEME; CHLOROPLAST; PHOTOSYNTHESIS; TRANSMEMBRANE.
FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND)
FT (BY SIMILARITY).
FT METAL 100 100 IRON 2 (HEME AXIAL LIGAND)
FT (BY SIMILARITY).
FT METAL 187 187 IRON 2 (HEME AXIAL LIGAND)
FT (BY SIMILARITY).
FT METAL 202 202 IRON 1 (HEME AXIAL LIGAND)
FT (BY SIMILARITY).
SQ SEQUENCE 215 AA; 24164 MW; E18A90C1 CRC32;
Query Match 5.5%; Score 107; DB 1; Length 215;
Best Local Similarity 22.7%; Pred. No. 7.69e-02;
Matches 15; Conservative 21; Mismatches 24; Indels 6; Gaps 6;

Db 57 YIKPTVAEATSSVQIMNVNPGWLI-R-SI-HNWSAMVLMVLFVRLVLTGCFKRP 113
:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|
Qy 132 YIQMSIDNSILSVHTVSAQV-WNITRYDLYNFRPAVL-LMFLSVTKAFMETFVHL 189
:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|:::1:|
Db 114 REL-TW 118
:::1:|
Qy 190 CSLSGNSW 195
:::1:|

RESULT 9
ID CYB6_MOSSP STANDARD; PRT; 215 AA.
AC P12122;
DT 01-OCT-1989 (REL. 12, CREATED)
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
DE CYTOCHROME B6 (EC 1.10.99.1).
GN PETB.
OS NOSTOC SP. (STRAIN PCC 7906).
OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
OC CYANOBACTERIA (BLUE-GREEN ALGAE); NOSTOCALES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89008280.
RA KALLAS T., SPILLER S., MALKIN R.;
RL J. BIOL. CHEM. 263:14334-14342(1988).
CC -1- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART
CC OF THE CYANOBACTERIAL RESPIRATORY CHAIN.
CC -1- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCTANIN =
CC PLASTOQUINONE + 2 REDUCED PLASTOCTANIN.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.

```

RESULT 11
ID CYB6_MARPO STANDARD; PRT; 215 AA.
AC P06248;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE CYTOCHROME B6 (EC 1.10.99.1).
GN PETB
OS MARCHANTIA POLYTRORPHA (LIVERWORT).
OC CHLOROPLAST.
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; BRYOPHYTA; HEPATICOPSIDA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89068687.
RA FUKUZAWA H., KOHCHI T., SANO T., SHIRAI H., UMESONO K., INOKUCHI H.,
RA OZEKI H., OHYAMA K.
RL J. MOL. BIOL. 203:333-351(1988).
RN [2]
RP COMPLETE GENOME.
RA OHYAMA K., FUKUZAWA H., KOHCHI T., SHIRAI H., SANO T., SANO S.,
RA UMESONO K., SHIKI Y., TAKEUCHI M., CHANG E., ACTA S., INOKUCHI H.,
RA OZEKI H.;
RL NATURE 322:572-574(1986).
RN [3]
RP SEQUENCE FROM N.A.
RA FUKUZAWA H., YOSHIDA T., KOHCHI T., OKUMURA T., SAWANO Y., OHYAMA K.,
RL FEBS LETT. 220:61-65(1987).
CC -1- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART
CC OF THE CHLOROPLASTIC RESPIRATORY CHAIN.
CC -2- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PTOCANTANIN +
CC PLASTOQUINONE + 2 REDUCED PLASTOCYANIN.
CC -3- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -4- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,
CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.
CC -5- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.
DR EMBL: X04465; E122080; -.
DR PIR: A00163; CRLV6.
DR PIR: S01552; S01552.
DR PIR: S02432; S02432.
DR MENDEL: 2551; MARPO; PETB:1.
DR PROSITE: P800192; CYTOCHROME_B_HEME; 1.
KW ELECTRON TRANSPORT; HEME; CHLOROPLAST; PHOTOSYNTHESIS; TRANSMEMBRANE.
FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND)
FT (BY SIMILARITY)
FT METAL 100 100 IRON 2 (HEME AXIAL LIGAND)
FT (BY SIMILARITY)
FT METAL 187 187 IRON 2 (HEME AXIAL LIGAND)
FT (BY SIMILARITY)
FT METAL 202 202 IRON 1 (HEME AXIAL LIGAND)
FT (BY SIMILARITY)
SQ SEQUENCE 215 AA; 24307 MW; A8DE4A12 CRC32;

```


Sun Jan 24 11:29:03 1999

US-09-162-597-3.rsp

Page 7

FT METAL 107 107 IRON 2 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 FT METAL 194 194 IRON 2 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 FT METAL 209 209 IRON 1 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 SQ SEQUENCE 222 AA; 25179 MW; 7F5AB232 CRC32;

Query Match 5.44; Score 105; DB 1; Length 222;
 Best Local Similarity 21.24; Pred. No. 1.45e-01;
 Matches 14; Conservative 21; Mismatches 25; Indels 6; Gaps 6;

Db 64 YTKPTVAEFTSVQIMNEVNFGLI-R-SI-HRWSASMMVLMILHIFRVLTGGFKRP 120
 | : : : | : | : | : : : | : : : | : : : | : : :
 Qy 132 YIQMSIDNSISLVHIVASAOV-WMITRYDLYNFRPAVL-LLMFSLVTKAFVMTFVHL 189

Db 121 REL-TW 125

Qy 190 CSLGSW 195

RESULT 15
 ID CYB6_CTAPA STANDARD; PRT; 215 AA.
 AC P48121;
 DT 01-FEB-1996 (REL. 33, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE CYTOCHROME B6 (EC 1.10.99.1).
 GN PETB.
 OS CYANOPHORA PARADOXA.
 OG CYANELLE.
 OC EUKARYOTA; PLANTA; PHYCOPHYTA; GLAUCOPHYTA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-L8555 / PRINGSHEIM;
 RA STIREWALT V.L., MICHALOWSKI C.B., LUFFELHARDT W., BOHNERT H.J.,
 RA BRYANT D.A.;
 RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: COMPONENT OF THE CYTOCHROME B6/F COMPLEX WHICH IS PART
 CC OF THE CYANELLE RESPIRATORY CHAIN.
 CC -1- CATALYTIC ACTIVITY: PLASTOQUINOL-1 + 2 OXYDIZED PLASTOCYANIN =
 CC PLASTOQUINONE + 2 REDUCED PLASTOCYANIN.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN.
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,
 CC 17 KD POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.
 CC -1- SIMILARITY: WITH THE AMINO END OF MITOCHONDRIAL CYTOCHROME B.
 DR EMBL; U30821; G1016119; -.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW ELECTRON TRANSPORT; HEME; CYANELLE; PHOTOSYNTHESIS; TRANSMEMBRANE.
 FT METAL 86 86 IRON 1 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 FT METAL 100 100 IRON 2 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 FT METAL 187 187 IRON 2 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 FT METAL 202 202 IRON 1 (HEME AXIAL LIGAND)
 FT (BY SIMILARITY).
 SQ SEQUENCE 215 AA; 24188 MW; 1042B361 CRC32;

Query Match 5.34; Score 103; DB 1; Length 215;
 Best Local Similarity 23.74; Pred. No. 2.66e-01;
 Matches 18; Conservative 21; Mismatches 31; Indels 6; Gaps 6;

Db 57 YTRPTVTEAFASVQIMTVNFGLI-R-ST-HRWSASMMVLMILHIFRVLTGGFKRP 113
 | : : : | : | : | : : : | : : : | : : : | : : :
 Qy 132 YIQMSIDNSISLVHIVASAOV-WMITRYDLYNFRPAVL-LLMFSLVTKAFVMTFVHL 189

Db 114 REL-TWVGVILAVIT 128

Qy 190 CSLGSWRLDARAVVT 205

Search completed: Fri Jan 22 18:00:58 1999
 Job time: 18 secs.

(TM)

[illegible]

GENETICS

#gene	petB
-------	------

```

GENETICS
#gene      petB
#map_position  CP1232-71237, 72049-72690
#genome    chloroplast
#introns   2/3
CLASSIFICATION
#superfamily cytochrome b6; cytochrome b6 homology
KEYWORDS
#chloroplast; chromoprotein; electron transfer; heme; iron;
#oxidoreductase; photosynthesis; thylakoid
FEATURE
15-215      #domain cytochrome b6 homology |label CB6|
86,202      #binding_site heme iron (His (axial ligands) #status
            predicted|
100,187     #binding_site heme iron (His (axial ligands) #status
            predicted|
SUMMARY
#length 215 #molecular-weight 24182 #checksum 9304

```

Db 76 YY

50 76 11

```
#journal      Nucleic Acids Res. (1989) 17:2859-2860
#title..... Nucleotide sequence of the 5.2 kbp-barley chloroplast-DNA
               fragment, containing psbB-psbH-petB-petD gene cluster.
#cross-references MIMD:89240047
#accession    S04149
               #molecule_type DNA
               #residues      1-232  #label REV
               #cross-references EMBL:X14107; NID:q11593; PID:q11596

GENETICS
#gene         petB
#genome       chloroplast
CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology
KEYWORDS       alternative splicing; chloroplast; chromoprotein; electron
               transfer; heme; iron; oxidoreductase; photosynthesis;
               thylakoid

FEATURE
32-232         #domain cytochrome b6 homology #label Cb6\
103,204        #binding_site heme iron, low potential (His) (axial
               ligands) #status predicted\
117,219        #binding_site heme iron, high potential (His) (axial
               ligands) #status predicted\

SUMMARY
#length 232 #molecular_weight 26078 #checksum 4328

Query Match          5.8%; Score 113; DB 1; Length 232;
Best Local Similarity 23.5%; Pred. No. 8.32e-02;
Matches 20; Conservative 27; Mismatches 32; Indels 6; Gaps 6;

Db 74 TYRPTVEAFSSVOITMEANFGWLI-R-SV-HWNASMVMYLMILVFFVYLNGFFKPK 130
Qy 132 YIQMSIDNSILSVHIVASQAQ-WMITRDLYLHFNPAFL-LNLFSLVYKAFMETFVFL 189

Db 131 REL-TWTVGTVGLAVLTASFVGYSYS 154
Qy 190 CSLGSWALDRAVTVGLLAKRFG 214

RESULT 7
ENTRY
TITLE      S58581 #type complete
           plastocinonol--plastocyanin reductase (EC 1.10.99.1)
           cytochrome b6 - maize chloroplast
ORGANISM   #formal_name chloroplast Zea mays #common_name maize
DATE       29-Nov-1995 #sequence_revision 19-Jan-1996 #text_change
           08-Sep-1997
ACCESSIONS S58581
REFERENCE
#authors   Maier, R.M.; Neckermann, K.; Igloi, G.L.; Koessel, H.
#journal   J. Mol. Biol. (1995) 251:614-628
#title     Complete sequence of the maize chloroplast genome: gene
           content, hotspots of divergence and fine tuning of genetic
           information by transcript editing.
#accession S58581
#status    nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues  1-234  #label MAI
#cross-references EMBL:X86563; NID:g902200; PID:g902251
#note      the nucleotide sequence was submitted to the EMBL Data
           Library, April 1995

GENETICS
#gene         petB
#genome       chloroplast
#introns      2/3
CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology
KEYWORDS       chloroplast; electron transfer; heme; membrane-associated
               complex; oxidoreductase; photosynthesis; thylakoid

FEATURE
34-234        #domain cytochrome b6 homology #label Cb6
SUMMARY       #length 234 #molecular_weight 26182 #checksum 5835

Query Match          5.8%; Score 113; DB 2; Length 234;
Best Local Similarity 23.5%; Pred. No. 8.32e-02;
Matches 20; Conservative 27; Mismatches 32; Indels 6; Gaps 6;

Db 76 TYRPTVEAFSSVOITMEANFGWLI-R-SV-HWNASMVMYLMILVFFVYLNGFFKPK 132
```

Qy 132 YIQMSIDNSLIVHIVASAOV-WMITRDLTYNFRPAVL-LIMFLSYKAFVWETFYHL 189
 Db 133 REL-TWTGTVLAVLTASFGVTGYS 156
 Qy 190 CSLGSWALDARAVYGLLALHFG 214

RESULT 8
 ENTRY CBMT5 #type complete
 TITLE plastoquinol-plastocyanin reductase (EC 1.10.9.9.1)
 cytochrome b6 - common tobacco chloroplast
 ALTERNATE_NAMES cytochrome b563
 ORGANISM #fomal_name chloroplast Nicotiana tabacum #common_name common tobacco
 DATE 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 04-Oct-1996
 ACCESSIONS A00162
 REFERENCE A00149
 #authors Sugiura, M.
 #submission submitted to the EMBL Data Library, August 1986
 #accession A00162
 #molecule_type DNA
 #residues 1-215 #label SUG
 #experimental_source cv. Bright Yellow 4
 REFERENCE A38013
 #authors Shinozaki, K.; Ohme, M.; Tanaka, M.; Wakasugi, T.; Hayashida, N.; Matsubayashi, T.; Zalta, N.; Chunwongse, J.; Obokata, J.; Yamaguchi-Shinozaki, K.; Ohto, C.; Torazawa, K.; Meng, B.Y.; Sugita, M.; Denc, H.; Kanogashira, T.; Yamada, K.; Kusuda, J.; Takaiwa, F.; Kato, A.; Tohdoh, N.; Shimada, H.; Sugiura, M.
 #journal EMBO J. (1986) 5:2043-2049
 #title The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression.
 #contents annotation; gene organization, sites, features
 COMMENT This cytochrome is one of the components of a specific stoichiometric cytochrome b6-f complex that contains two molecules of cytochrome b6, one cytochrome f, and one nonheme iron-sulfur center.

GENETICS
 #gene petB
 #genome chloroplast
 #introns 2/3
 CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology
 KEYWORDS chloroplast; chromoprotein; electron transfer; heme; iron; oxidoreductase; photosynthesis; thylakoid

FEATURE
 15-215 #domain cytochrome b6 homology #label CB6
 86,202 #binding_site heme iron (His) (axial ligands) #status predicted
 100,187 #binding_site heme iron (His) (axial ligands) #status predicted
 SUMMARY #length 215 #molecular-weight 24136 #checksum 9633

Query Match 5.7%; Score 111; DB 1; Length 215;
 Best Local Similarity 22.4%; Pred. No. 1.44e-01;
 Matches 19; Conservative 28; Mismatches 32; Indels 6; Gaps 6;

Db 57 YIYRTVFETASFOQIMTEANFWGLI-R-SV-HRWSAMWMLVHLFVRYTLGGFKPK 113
 Qy 132 YIQMSIDNSLIVHIVASAOV-WMITRDLTYNFRPAVL-LIMFLSYKAFVWETFYHL 189
 Db 114 REL-TWTGTVLAVLTASFGVTGYS 137
 Qy 190 CSLGSWALDARAVYGLLALHFG 214

RESULT 9
 ENTRY S52970 #type complete
 TITLE NADH dehydrogenase (ubiquinone) (BC 1.6.5.3) chain 6 - honeybee mitochondrion (SGC4)
 ORGANISM #fomal_name mitochondrion Apis mellifera #common_name

honeybee
 DATE 15-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 21-Aug-1998
 ACCESSIONS S52970
 REFERENCE S52960
 #authors Crozier, R.H.; Crozier, Y.C.
 #journal Genetics (1993) 133:97-117
 #title The mitochondrial genome of the honeybee Apis mellifera: complete sequence and genome organization.
 #accession S52970
 #molecule_type DNA
 #residues 1-167 #label CRO
 #cross-references EMBL:L06178; NID:g336279; PID:g829009
 #experimental_source ligustica

GENETICS
 #genome mitochondrion
 #genetic_code SGC4
 #start_codon ATT
 CLASSIFICATION #superfamily NADH dehydrogenase (ubiquinone) chain 6
 KEYWORDS membrane-associated complex; mitochondrion; NAD; oxidoreductase
 SUMMARY #length 167 #molecular-weight 19971 #checksum 6545

Query Match 5.5%; Score 106; DB 2; Length 167;
 Best Local Similarity 29.9%; Pred. No. 5.58e-01;
 Matches 20; Conservative 17; Mismatches 25; Indels 5; Gaps 5;

Db 11 IFMSSLSIMLITY-LNNIFNSPMLIYLYSYTMSLMPT-MCSMISLLIMLITVF 68
 Qy 147 IVASAOVWMI-TRITLYNFRPAVLMLPLSYKAFV-METVHFLCSLGSWALDARAV- 203
 Db 69 LSGMLM 75
 Qy 204 VTGLLAL 210

RESULT 10
 ENTRY S21253 #type complete
 TITLE plastoquinol-plastocyanin reductase (EC 1.10.9.9.1)
 cytochrome b6 - Chlamydomonas reinhardtii chloroplast
 ORGANISM #fomal_name chloroplast Chlamydomonas reinhardtii
 DATE 03-Feb-1994 #sequence_revision 02-Aug-1994 #text_change 05-Sep-1997
 ACCESSIONS S21253; S16917
 REFERENCE S20938
 #authors Huang, C.; Liu, X.Q.
 #journal Plant Mol. Biol. (1992) 18:985-988
 #title Nucleotide sequence of the frx, petB and trnL genes in the chloroplast genome of Chlamydomonas reinhardtii.
 #cross-references MUID:92256821
 #accession S21253
 #status translation not shown
 #molecule_type DNA
 #residues 1-215 #label HUA
 #cross-references EMBL:X62905; NID:g12497; PID:g12499

REFERENCE S16916
 #authors Bueschlen, S.; Choquet, Y.; Kuras, R.; Wollman, P.A.
 #journal FEBS Lett. (1991) 284:257-262
 #title Nucleotide sequences of the continuous and separated petA, petB and petD chloroplast genes in Chlamydomonas reinhardtii.
 #cross-references MUID:91285146
 #accession S16917
 #molecule_type DNA
 #residues 1-215 #label BUE
 #cross-references EMBL:X72918; NID:g603530; PID:g288909

GENETICS
 #gene petB
 #genome chloroplast
 CLASSIFICATION #superfamily cytochrome b6; cytochrome b6 homology
 KEYWORDS chloroplast; chromoprotein; electron transfer; heme; iron; oxidoreductase; photosynthesis; thylakoid

FEATURE

```
ENTRY      B69031      #type complete
TITLE      conserved hypothetical protein MTH129 - Methanobacterium
            thermoautotrophicum (strain Delta H)
ORGANISM    #formal_name Methanobacterium thermoautotrophicum
DATE        05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
            13-Sep-1998
ACCESSIONS  B69031
REFERENCE   A69000
            #authors Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.;
```

Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.; Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.; Lum, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Caruso, A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.; McDougall, S.; Shimer, G.; Goyal, A.; Pietrovski, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Moelling, J.; Reeve, J.N.

#journal J. Bacteriol. (1997) 179:7135-7155

#title Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional analysis and comparative genomics.

#cross-references MUID:98037514

#accession B69031

#status preliminary; nucleic acid sequence not shown; translation not shown

#molecule_type DNA

#residues 1-212 #label MTH

#cross-references GB:AE000890; GB:AE000666; NID:g2622331; PID:g2622340

#experimental_source strain Delta H

GENETICS

#gene MTH1229

CLASSIFICATION

#superfamily Methanococcus jannaschii conserved hypothetical protein MJ0805

SUMMARY

#length 212 #molecular-weight 23956 #checksum 2161

Query Match 5.4%; Score 105; DB 2; Length 212;

Best Local Similarity 31.5%; Pred. No. 7.28e-01;

Matches 23; Conservative 17; Mismatches 27; Indels 6; Gaps 6;

Db 44 ANWSELGIGSSIIITWIGPMQVKNIVDYDELEIRSEEVLFETVEFDE-OPSLRL 101

Qy 105 GWATLAL-INSRCIPLWVGARCI-EF-DWKYQMSIDSNISLVHYIVASQAQWMITRYDL 161

Db 102 AYERQIRILWMLLW 114

Qy 162 -YENFPAVLLW 173

RESULT 14

ENTRY

TITLE

CBLV6 #type complete

plastoquinol-plastocyanin reductase (EC 1.10.99.1)

cytochrome b6 - liverwort (Marchantia polymorpha)

chloroplast

ORGANISM

#formal_name chloroplast Marchantia polymorpha

DATE

30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 04-Oct-1996

ACCESSIONS

REFERENCE

S01552; S02432; A00163

S01529

#authors Fukuzawa, H.; Kohchi, T.; Sano, T.; Shirai, H.; Umesono, K.; Inokuchi, H.; Ozeki, H.; Ohya, K.

#journal J. Mol. Biol. (1988) 203:333-351

#title Structure and organization of Marchantia polymorpha chloroplast genome. III. Gene organization of the large single copy region from rbcL to trnI(CAO).

#cross-references MUID:8906867

#accession S01552

#molecule_type DNA

#residues 1-215 #label FOX

#cross-references EMBL:X04465

REFERENCE

A38014

#authors Ohya, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, T.; Sano, S.; Umesono, K.; Shiki, Y.; Takeuchi, M.; Chang, Z.; Aota, S.; Inokuchi, H.; Ozeki, H.

#journal Nature (1986) 322:572-574

#title Chloroplast gene organization deduced from complete sequence of liverwort Marchantia polymorpha chloroplast DNA. annotation; gene organization, sites, features

#contents

REFERENCE

S02432

#authors Fukuzawa, H.; Yoshida, T.; Kohchi, T.; Okumura, T.; Sawano, Y.; Ohya, K.

#journal FEBS Lett. (1987) 220:61-66

#title Splicing of group II introns in mRNAs coding for cytochrome

b6 and subunit IV in the liverwort Marchantia polymorpha chloroplast genome. Exon specifying a region coding for two genes with the spacer region.

#accession S02432

#status not compared with conceptual translation

#molecule_type mRNA

#residues 1-8 #label FUN

GENETICS

#gene petB

#genome chloroplast

#introns 2/3

CLASSIFICATION

KEYWORDS

#superfamily cytochrome b6; cytochrome b6 homology chloroplast; chromoprotein; electron transfer; heme; iron; oxidoreductase; photosynthesis; thylakoid

FEATURE

15-215 #domain cytochrome b6 homology #label CB6

86,187 #binding_site heme iron, low potential (His) (axial ligands) #status predicted

100,202 #binding_site heme iron, high potential (His) (axial ligands) #status predicted

SUMMARY

#length 215 #molecular-weight 24307 #checksum 8453

Query Match 5.4%; Score 105; DB 1; Length 215;

Best Local Similarity 22.7%; Pred. No. 7.28e-01;

Matches 15; Conservative 21; Mismatches 24; Indels 6; Gaps 6;

Db 57 YTRPTVTEAFSSVQYIMTEVNGWLI-R-SV-HRNSASMMVLMILHIFRVYLTGCFKXP 113

Qy 132 YIQMSIDSNISLVHYIVASQAQ-WMITRYDLHNFPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TW 118

Qy 190 CSLGSG 195

RESULT 15

ENTRY

TITLE

cytochrome b6 - red alga (Porphyra purpurea) chloroplast

ORGANISM

#formal_name chloroplast Porphyra purpurea

DATE

19-Mar-1997 #sequence_revision 09-May-1997 #text_change 24-Apr-1998

ACCESSIONS

REFERENCE

S73262

S73108

#authors Reith, M.; Munnholland, J.

#journal Plant Mol. Biol. Rep. (1995) 13:333-335

#title Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.

#accession S73262

#status preliminary; nucleic acid sequence not shown; translation not shown

#molecule_type DNA

#residues 1-215 #label REI

#cross-references EMBL:X38804; NID:g1276652; PID:g1276807

#note the nucleotide sequence was submitted to the EMBL Data Library, October 1995

GENETICS

#gene petB

#genome chloroplast

CLASSIFICATION

KEYWORDS

#superfamily cytochrome b6; cytochrome b6 homology chloroplast; thylakoid

FEATURE

15-215 #domain cytochrome b6 homology #label CB6

SUMMARY

#length 215 #molecular-weight 24224 #checksum 9699

Query Match 5.4%; Score 104; DB 2; Length 215;

Best Local Similarity 22.7%; Pred. No. 9.48e-01;

Matches 15; Conservative 20; Mismatches 25; Indels 6; Gaps 6;

Db 57 YTRPTVTEAFSSVQYIMTEVNGWLI-R-SV-HRNSASMMVLMILHIFRVYLTGCFKXP 113

Qy 132 YIQMSIDSNISLVHYIVASQAQ-WMITRYDLHNFPAVL-LLMFLSVYKAFVMTFVHL 189

Db 114 REL-TW 118

Sun Jan 24 11:29:02 1999

US-09-162-597-3.rpr

Page 7

Cy 190 CSLGSN 195

Search completed: Fri Jan 22 18:00:22 1999
Job time : 19 secs.

M P S R C H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Jan 22 18:03:07 1999; MPsrch time 3.58 Seconds
597.255 Million cell updates/sec
Tabular output not generated.

Title: >US-09-162-597-3
Description: (1-245) from US09162597.pgp
Perfect Score: 1940
Sequence: 1 MHLFPGMCFALAYFPFIT.....GLVSTLMTLPASLQVLVK 245

Scoring table: PAM 150
Gap 11

Searched: 92929 seqs, 8738560 residues

Post-processing: Minimum Match 04
Listing first 45 summaries

Database: a-issued
1:5.COMB 2:PCT9.COMB 3:backfiles1

Statistics: Mean 31.166; Variance 148.701; scale 0.210

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	No.	Score	Match	Length	ID	Description	Pred. No.
1	94	4.8	772	2	PCT-US95-0		Sequence 2, Applicatio	1.66e+01
2	94	4.8	772	1	US-08-258-		Sequence 2, Applicatio	1.66e+01
3	89	4.6	349	1	US-08-118-		Sequence 71, Applicati	3.57e+01
4	89	4.6	349	2	PCT-US93-0		Sequence 71, Applicati	3.57e+01
5	87	4.5	307	1	US-08-948-		Sequence 3, Applicatio	4.82e+01
6	87	4.5	417	1	US-08-553-		Sequence 2, Applicatio	4.82e+01
7	85	4.4	201	3	5489533-3		Patent No. 5489533.	6.50e+01
8	85	4.4	252	2	PCT-US95-1		Sequence 103, Applicat	6.50e+01
9	85	4.4	274	3	5489533-2		Patent No. 5489533.	6.50e+01
10	85	4.4	274	3	5512660-2		Patent No. 5512660.	6.50e+01
11	84	4.3	316	1	US-08-118-		Sequence 46, Applicati	7.54e+01
12	84	4.3	316	2	PCT-US93-0		Sequence 46, Applicati	7.54e+01
13	83	4.3	346	2	PCT-US96-1		Sequence 2, Applicatio	8.75e+01
14	83	4.3	363	1	US-08-148-		Sequence 2, Applicatio	8.75e+01
15	83	4.3	363	1	US-08-148-		Sequence 3, Applicatio	8.75e+01
16	83	4.3	380	2	PCT-US93-1		Sequence 40, Applicati	8.75e+01
17	83	4.3	380	1	US-08-153-		Sequence 40, Applicati	8.75e+01
18	84	4.3	494	2	PCT-US94-0		Sequence 4, Applicatio	7.54e+01
19	84	4.3	494	1	US-08-464-		Sequence 4, Applicatio	7.54e+01
20	84	4.3	2368	1	US-08-198-		Sequence 15, Applicati	7.54e+01
21	82	4.2	225	3	5436139-4		Patent No. 5436139.	1.01e+02
22	82	4.2	226	3	5436139-5		Patent No. 5436139.	1.01e+02
23	82	4.2	226	3	5196194-21		Patent No. 5196194.	1.01e+02

24	82	4.2	226	2	PCT-US96-1	Sequence 14, Applicati	1.01e+02
25	82	4.2	226	3	5198348-1	Patent No. 5198348.	1.01e+02
26	82	4.2	226	1	US-08-378-	Sequence 3, Applicatio	1.01e+02
27	82	4.2	236	1	US-08-378-	Sequence 1, Applicatio	1.01e+02
28	82	4.2	277	3	5164485-2	Patent No. 5164485.	1.01e+02
29	82	4.2	279	1	US-07-688-	Sequence 30, Applicati	1.01e+02
30	82	4.2	279	2	PCT-US91-0	Sequence 29, Applicati	1.01e+02
31	82	4.2	281	1	US-08-458-	Sequence 214, Applicat	1.01e+02
32	82	4.2	281	1	US-08-105-	Sequence 214, Applicat	1.01e+02
33	82	4.2	351	2	PCT-US96-1	Sequence 4, Applicatio	1.01e+02
34	82	4.2	361	2	PCT-US93-0	Sequence 4, Applicatio	1.01e+02
35	82	4.2	389	1	US-08-709-	Sequence 216, Applicat	1.01e+02
36	82	4.2	389	1	US-08-709-	Sequence 219, Applicat	1.01e+02
37	82	4.2	389	1	US-08-105-	Sequence 219, Applicat	1.01e+02
38	82	4.2	389	1	US-08-458-	Sequence 216, Applicat	1.01e+02
39	82	4.2	389	1	US-08-458-	Sequence 219, Applicat	1.01e+02
40	82	4.2	389	1	US-08-105-	Sequence 216, Applicat	1.01e+02
41	82	4.2	395	3	5196194-18	Patent No. 5196194.	1.01e+02
42	82	4.2	397	2	PCT-US96-1	Sequence 6, Applicatio	1.01e+02
43	82	4.2	502	1	US-08-496-	Sequence 4, Applicatio	1.01e+02
44	82	4.2	502	1	US-08-466-	Sequence 10, Applicati	1.01e+02
45	82	4.2	995	2	PCT-US93-0	Sequence 14, Applicati	1.01e+02

ALIGNMENTS

RESULT 1
ID PCT-US95-07391A-2 STANDARD; PRT; 772 AA.
XX
AC XXXXXX
XX
XX
XX
XX
DE Sequence 2, Application PC/TUS9507391A
XX
CC Sequence 2, Application PC/TUS9507391A
CC GENERAL INFORMATION:
CC APPLICANT: IBEX TECHNOLOGIES and
CC APPLICANT: ZIMMERMANN, Joseph
CC TITLE OF INVENTION: Nucleic Acid Sequences And Expression
CC TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived Fr
CC TITLE OF INVENTION: Flavobacterium heparinum
CC NUMBER OF SEQUENCES: 26
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hale and Dorr
CC STREET: 1455 Pennsylvania Avenue, N.W.
CC CITY: Washington, D.C.
CC COUNTRY: U.S.A.
CC ZIP: 20004
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/07391A
CC FILING DATE: 09-JUNE-1995
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/258,639
CC FILING DATE: 10 JUNE 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BAKER, Hollie L.
CC REGISTRATION NUMBER: 31,321
CC REFERENCE/DOCKET NUMBER: 104385.116PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)942-8400
CC TELEFAX: (202)942-8484
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 772 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear

Sun Jan 24 11:29:01 1999

US-09-162-597-3.ra1

Page 2

CC MOLECULE TYPE: protein
SQ SEQUENCE 772 AA; 87638 MW; 3109034 CN;

Query Match 4.8%; Score 94; DB 2; Length 772;
Best Local Similarity 26.1%; Pred. No. 1.66e+01;
Matches 23; Conservative 25; Mismatches 35; Indels 5; Gaps 3;

Db 94 DFFRYFNQGLTVRVELMALNYLMTKDPKVGREAITSIIDLETATFKPAGDISRGIGLF 153

Qy 66 DFIFGEFKASVDV-ADLIGLNLVMSRNA--GKGKYIMVAALGWAT-AELINSCRIPLM 120

Db 154 MYTGAIVYDWCYDQLKPEEKTRFKAFV 181

Qy 121 VGARGIEFDWKYIQMSIDSNISLVHYIV 148

RESULT 2

ID US-08-258-639A-2 STANDARD; PRT; 772 AA.

XX

AC XXXXXX

XX

DT

XX

DE Sequence 2, Application US/08258639A

XX

CC Sequence 2, Application US/08258639A

CC Patent No. 5681733

CC GENERAL INFORMATION:

CC APPLICANT: Su, Hongsheng

CC APPLICANT: Blain, Francoise

CC APPLICANT: Bennett, Clark

CC APPLICANT: Gu, Kangfu

CC APPLICANT: Zimmermann, Joseph

CC APPLICANT: Musil, Roy

CC TITLE OF INVENTION: Nucleic Acid Sequences And Expression

CC TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From

CC TITLE OF INVENTION: Flavobacterium heparinum

CC NUMBER OF SEQUENCES: 26

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Hale and Dorr

CC STREET: 1455 Pennsylvania Avenue, N.W.

CC CITY: Washington, D.C.

CC COUNTRY: U.S.A.

CC ZIP: 20004

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/258,639A

CC FILING DATE: 10 JUNE 1994

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Healey, William J.

CC REGISTRATION NUMBER: 36,160

CC REFERENCE/DOCKET NUMBER: 104385.116

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (202)942-8400

CC TELEFAX: (202)942-8484

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 772 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

SQ SEQUENCE 772 AA; 87638 MW; 3109034 CN;

Query Match 4.8%; Score 94; DB 1; Length 772;
Best Local Similarity 26.1%; Pred. No. 1.66e+01;
Matches 23; Conservative 25; Mismatches 35; Indels 5; Gaps 3;

Db 94 DFFRYFNQGLTVRVELMALNYLMTKDPKVGREAITSIIDLETATFKPAGDISRGIGLF 153

Qy 66 DFIFGEFKASVDV-ADLIGLNLVMSRNA--GKGKYIMVAALGWAT-AELINSCRIPLM 120

Db 154 MYTGAIVYDWCYDQLKPEEKTRFKAFV 181

Qy 121 VGARGIEFDWKYIQMSIDSNISLVHYIV 148

RESULT 3

ID US-08-118-270-71 STANDARD; PRT; 349 AA.

XX

AC XXXXXX

XX

DT

XX

DE Sequence 71, Application US/08118270

XX

CC Sequence 71, Application US/08118270

CC Patent No. 5508384

CC GENERAL INFORMATION:

CC APPLICANT: Murphy, Randall B.

CC APPLICANT: Schuster, David I.

CC TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN

CC TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF

CC NUMBER OF SEQUENCES: 348

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: BROWDY AND NEIMARK

CC STREET: 419 Seventh Street, N.W., Suite 300

CC CITY: Washington

CC STATE: D.C.

CC COUNTRY: USA

CC ZIP: 20004

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/118,270

CC FILING DATE: 09-SEP-1993

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 07/943,236

CC FILING DATE: 10-SEP-1992

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Townsend, Kevin G.

CC REGISTRATION NUMBER: 34,033

CC REFERENCE/DOCKET NUMBER: MURPHY-2A

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 202-626-5197

CC TELEFAX: 202-737-3528

CC TELEX: 248633

CC INFORMATION FOR SEQ ID NO: 71:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 349 amino acids

CC TYPE: amino acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: peptide

SQ SEQUENCE 349 AA; 39913 MW; 695896 CN;

Query Match 4.6%; Score 89; DB 1; Length 349;
Best Local Similarity 20.5%; Pred. No. 3.57e+01;
Matches 18; Conservative 29; Mismatches 38; Indels 3; Gaps 3;

Db 94 AIVADRHQVIMPLAPRISITKGVIIIVWVFFSLPFAICQKLFYKSIDIVRSCL 153

Qy 125 GIEFQYTIQMSIDSNISLVHYIVASQWMTIRYDLYNFRPAAVLMFL-IVTKAFVW 183

Db 154 DPPEFPAOL-PWKYLDIATFILLYLLPL 180

Qy 184 ETVFHLCSLGSWALDARA-VVTGLIAL 210


```
AC      xxxxxx
XX
DT
XX
DE      Sequence 3, Application US/08948616
XX
CC      Sequence 3, Application US/08948616
```

Query Match 4.5%; Score 87; DB 1; Length 307;
Best Local Similarity 26.9%; Pred. No. 4.82e+01;
Matches 14; Conservative 14; Mismatches 22; Indels 2; Gaps 2;

Db 240 YQCYLLVYYTGW-VRNVKSLTFGLICLNK-YLYELRNLQLFFHVTVGAFV 289
| | : | : | : | : | : | : | : | :
Qy 21 YKCSGLSEYNAFWCKVQAGVTYLFVLCKMLFLATFFPTWEGGIYDFIGEFM 72

RESULT 6
ID US-08-553-703A-2 STANDARD: PRT: 417 AA

AC XXXXXX
XT
DT
DE Sequence 2, Application US/08553703A
XX
CC Sequence 2, Application US/08553703A
CC Patent No. 5795767
CC GENERAL INFORMATION:
CC APPLICANT: MARU, ISAFUMI
CC APPLICANT: OHTA, YASUHIRO
CC APPLICANT: TSUKADA, YOJI
CC TITLE OF INVENTION: EPIMERASE
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/118,270
CC FILING DATE: 09-SEP-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/943,236
CC FILING DATE: 10-SEP-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Townsend, Kevin G.

CC REGISTRATION NUMBER: 34,033
CC REFERENCE/DOCKET NUMBER: MURPHY-2A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-628-5197
CC TELEFAX: 202-737-3528
CC TELEX: 248633
CC INFORMATION FOR SEQ ID NO: 46:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 316 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
SQ SEQUENCE 316 AA; 36040 MW; 553094 CN;

Query Match 4.3%; Score 84; DB 1; Length 316;
Best Local Similarity 37.0%; Pred. No. 7.54e+01;
Matches 10; Conservative 10; Mismatches 5; Indels 2; Gaps 2;

Db 43 ASADLIIA-CGLPFWATTIANNFWLWF 68
Qy 107 ATAEILMSRC-IPLVWGARGIEFDWKY 132

RESULT 12
ID PCT-US93-08528-46 STANDARD; PRT: 316 AA.
XX
AC XXXXXX
XX
DT
XX
DE Sequence 46, Application PC/TUS9308528
XX
CC Sequence 46, Application PC/TUS9308528
CC GENERAL INFORMATION:
CC APPLICANT: New York University
CC TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
CC TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
CC NUMBER OF SEQUENCES: 348
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: BROWDY AND WEINMARK
CC STREET: 419 Seventh Street, N.W., Suite 300
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20004
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/08528
CC FILING DATE: 09-SEP-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/943,236
CC FILING DATE: 10-SEP-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Townsend, Kevin G.
CC REGISTRATION NUMBER: 34,033
CC REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-628-5197
CC TELEFAX: 202-737-3528
CC TELEX: 248633
CC INFORMATION FOR SEQ ID NO: 46:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 316 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
SQ SEQUENCE 316 AA; 36040 MW; 553094 CN;

Query Match 4.3%; Score 84; DB 2; Length 316;
Best Local Similarity 37.0%; Pred. No. 7.54e+01;
Matches 10; Conservative 10; Mismatches 5; Indels 2; Gaps 2;

Db 43 ASADLIIA-CGLPFWATTIANNFWLWF 68
Qy 107 ATAEILMSRC-IPLVWGARGIEFDWKY 132

RESULT 13
ID PCT-US96-10602-2 STANDARD; PRT: 346 AA.
XX
AC XXXXXX
XX
DT
XX
DE Sequence 2, Application PC/TUS9610602
XX
CC Sequence 2, Application PC/TUS9610602
CC GENERAL INFORMATION:
CC APPLICANT: The General Hospital Corporation
CC TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/10602
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 60/017,814
CC FILING DATE: 20-JUN-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Clark, Paul T.
CC REGISTRATION NUMBER: 30,162
CC REFERENCE/DOCKET NUMBER: 00786/282001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617/542-5070
CC TELEFAX: 617/542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 346 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 346 AA; 40128 MW; 678897 CN;

Query Match 4.3%; Score 83; DB 2; Length 346;
Best Local Similarity 35.1%; Pred. No. 8.75e+01;
Matches 13; Conservative 10; Mismatches 10; Indels 4; Gaps 4;

Db 192 NGFNMVLRRIIVLVLLC-LIFLVLL-DWGLI 226
Qy 30 AP-WVCVQAGVTLFVQL-CXMLELAFVFFVWGGI 64

RESULT 14
ID US-08-148-209A-2 STANDARD; PRT: 363 AA.
XX
AC XXXXXX

XX
DT
XX
DE Sequence 2, Application US/08148209A
XX
CC Sequence 2, Application US/08148209A
CC Patent No. 5556780
CC GENERAL INFORMATION:
CC APPLICANT: Dzaou, Victor J
CC APPLICANT: Mukoyama, Masashi
CC TITLE OF INVENTION: TYPE-2 ANGIOTENSIN II RECEPTOR AND GENE
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
CC STREET: 4 Embarcadero Center, Suite 3400
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-4187
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/148,209A
CC FILING DATE: 05-NOV-1993
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Rowland, Bertram I
CC REGISTRATION NUMBER: 20,015
CC REFERENCE/DOCKET NUMBER: A-58491-1/BIR
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 781-1989
CC TELEFAX: (415) 398-3249
CC TELEX: 910 277299
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 363 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 363 AA: 41373 MW; 722740 CN;

Query Match 4.3%; Score 83; DB 1; Length 363;
Best Local Similarity 32.6%; Pred. No. 8.75e+01;
Matches 14; Conservative 11; Mismatches 14; Indels 4; Gaps 4;

Db 172 CLSSLPITY-FRD-VRI-IEYLGVNACIMAPPPEKYAQSAGI 211
| |: : | : |: : || | | : : |: : ||
Qy 23 C-SGLSEYNAPKWCQAGVITLFLVOLCKMLPLATFFPTWEGGI 64

RESULT 15
ID US-08-148-209A-3 STANDARD; PRT; 363 AA.

XX
AC *****
XX
DT
XX
DE Sequence 3, Application US/08148209A
XX
CC Sequence 3, Application US/08148209A
CC Patent No. 5556780
CC GENERAL INFORMATION:
CC APPLICANT: Dzaou, Victor J
CC APPLICANT: Mukoyama, Masashi
CC TITLE OF INVENTION: TYPE-2 ANGIOTENSIN II RECEPTOR AND GENE
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
CC STREET: 4 Embarcadero Center, Suite 3400
CC CITY: San Francisco

CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-4187
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/148,209A
CC FILING DATE: 05-NOV-1993
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Rowland, Bertram I
CC REGISTRATION NUMBER: 20,015
CC REFERENCE/DOCKET NUMBER: A-58491-1/BIR
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 781-1989
CC TELEFAX: (415) 398-3249
CC TELEX: 910 277299
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 363 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
SQ SEQUENCE 363 AA: 41350 MW; 725910 CN;

Query Match 4.3%; Score 83; DB 1; Length 363;
Best Local Similarity 32.6%; Pred. No. 8.75e+01;
Matches 14; Conservative 11; Mismatches 14; Indels 4; Gaps 4;

Db 172 CLSSLPITY-FRD-VRI-IEYLGVNACIMAPPPEKYAQSAGI 211
| |: : | : |: : || | | : : |: : ||
Qy 23 C-SGLSEYNAPKWCQAGVITLFLVOLCKMLPLATFFPTWEGGI 64

Search completed: Fri Jan 22 18:03:21 1999
Job time : 14 secs.

 M P S R C H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Jan 22 17:56:21 1999; MasPar time 8.04 Seconds
 492.876 Million cell updates/sec

Tabular output not generated.

Title: >US-09-162-597-3
 Description: (1-245) from US09162597.peg
 Perfect Score: 1940
 Sequence: 1 MTLFHPGNCFLALYPPYFIT.....GLVSQTLMLYFASLQVLVK 245

Scoring table: PAM 150
 Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 04
 Listing first 45 summaries

Database: a:geneseq32
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29

Statistics: Mean 33.119; Variance 147.508; scale 0.225

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	99	5.1	237	22	W20142	H. pylori inner membr	1.14e+01
2	98	5.1	422	21	R97722	Mouse inositol polyph	1.34e+01
3	98	5.1	497	22	W20821	H. pylori cell envelo	1.34e+01
4	95	4.9	1353	19	R99251	Murine adenylate cycl	2.17e+01
5	94	4.8	772	15	R89264	Heparinase-II.	2.55e+01
6	94	4.8	843	13	R67760	Lys-aminopeptidase Pe	2.55e+01
7	91	4.7	359	22	W20927	H. pylori surface or	4.11e+01
8	92	4.7	521	21	W14445	CarA gene product.	3.51e+01
9	89	4.6	349	19	W02722	G-protein coupled nou	5.64e+01
10	89	4.6	349	16	R48750	G-protein coupled nou	5.64e+01
11	87	4.5	174	3	R21412	NADH dehydrogenase 6	7.70e+01
12	87	4.5	293	21	W14481	Ramp-1.	7.70e+01
13	87	4.5	417	16	R79929	Porcine acylglucosami	7.70e+01
14	88	4.5	558	2	P70429	Murine neuroleukin.	6.59e+01
15	85	4.4	152	29	W48271	Rat anjurin 1.	1.05e+02
16	85	4.4	252	20	W01220	Serine protease PfSP1	1.05e+02
17	85	4.4	274	1	R06836	Intercellular adhesio	1.05e+02
18	85	4.4	1477	22	W10424	Saccharomyces cerevis	1.05e+02

19	85	4.4	1477	13	R67691	S. cerevisiae scaur2R	1.05e+02
20	86	4.4	1989	18	R99640	Peripheral-nervous-sy	8.99e+01
21	86	4.4	4473	20	R97244	Virulence gene cluste	8.99e+01
22	84	4.3	236	1	P81992	Hepatitis B viral sur	1.22e+02
23	84	4.3	316	19	W02697	G-protein coupled rat	1.22e+02
24	84	4.3	316	16	R48725	G-protein coupled rat	1.22e+02
25	83	4.3	346	21	W09046	WHV core-surface fusi	1.42e+02
26	83	4.3	363	13	R66934	Mouse AT2 receptor.	1.42e+02
27	83	4.3	380	11	R53750	Seven transmembrane r	1.42e+02
28	83	4.3	391	29	W38453	Wild-type human homol	1.42e+02
29	84	4.3	422	22	W20249	H. pylori transmembra	1.22e+02
30	84	4.3	426	22	W20934	H. pylori surface or	1.22e+02
31	84	4.3	494	28	W42996	Putative mature potas	1.22e+02
32	84	4.3	494	16	R07065	Human K+ channel 2 ma	1.22e+02
33	84	4.3	1233	29	W34536	Nudaurelia beta virus	1.22e+02
34	84	4.3	1233	29	W41935	Nudaurelia beta-like	1.22e+02
35	84	4.3	2368	26	W26663	Yeast checkpoint cont	1.22e+02
36	84	4.3	2510	6	R29527	HCV antigen 77N1-30.	1.22e+02
37	84	4.3	3011	4	R22154	NANBV Hutch c59 isola	1.22e+02
38	83	4.3	3085	23	W19701	ATM mutant G5170C.	1.42e+02
39	82	4.2	226	7	R33252	HBSag encoded by pGPD	1.66e+02
40	82	4.2	250	2	R11496	RF142/HBSag.	1.66e+02
41	82	4.2	351	21	W09047	Plasmid pB8V DN inser	1.66e+02
42	82	4.2	558	2	P70430	Human neuroleukin.	1.66e+02
43	82	4.2	658	27	W27666	Streptococcus pneumo	1.66e+02
44	82	4.2	724	27	W36797	Novel human gene, des	1.66e+02
45	82	4.2	898	21	W14777	Granulosis virus infe	1.66e+02

ALIGNMENTS

RESULT 1
 ID W20142 standard; protein; 237 AA.
 AC W20142;
 DT 08-JUL-1997 (first entry)
 DE H. pylori inner membrane protein, 14455461.aa.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 NW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
 OS Helicobacter pylori.
 FH Key Location/Qualifiers
 FT misc_difference 3
 FT /note= "encoded by YCC"
 FT misc_difference 130
 FT /note= "encoded by RCG"
 FT misc_difference 183
 FT /note= "encoded by GST"
 PN W09640893-A1.
 PD 13-DEC-1996.
 PF 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB.
 PI Berglindh OT, Smith D, Møllgaard BL;
 DR WPI; 97-052306/05.
 DR N-PSDB; 767385.
 PT Helicobacter pylori nucleic acid sequences and related
 PT peptide(s) - useful for vaccines to treat or prevent H. pylori
 PT infection, and to detect Helicobacter
 PS Claim 56; Pages 365; 1481pp; English.
 CC The present sequence is a Helicobacter pylori inner membrane protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 CC and the predicted coding regions defined by computer evaluation. To
 CC identify likely H. pylori antigens for vaccine development, the amino
 CC acid sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide

KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;

Fl modified site 04

FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 187
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 200
FT /label= CAM_kinase-II
FT /note= 'putative CAM kinase II phosphorylation
FT site'
FT modified_site 201
FT /label= A-kinase
FT /note= 'putative protein kinase A phosphorylation
FT site'
FT modified_site 261
FT /label= Casein_kinase-II
FT /note= 'putative casein kinase II phosphorylation
FT site'
FT modified_site 266
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 303
FT /label= CAM_kinase-II
FT /note= 'putative CAM kinase II phosphorylation
FT site'
FT modified_site 304
FT /label= A-kinase
FT /note= 'putative protein kinase A phosphorylation
FT site'
FT modified_site 356
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 358
FT /label= CAM_kinase-II
FT /note= 'putative CAM kinase II phosphorylation
FT site'
FT modified_site 370
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 432
FT /label= Casein_kinase-II
FT /note= 'putative casein kinase II phosphorylation
FT site'
FT modified_site 436
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT domain 503..610
FT /label= Calcineurin_binding_domain
FT modified_site 581
FT /label= Casein_kinase-II
FT /note= 'putative casein kinase II phosphorylation
FT site'
FT modified_site 588
FT /label= Casein_kinase-II
FT /note= 'putative casein kinase II phosphorylation
FT site'
FT modified_site 596
FT /label= CAM_kinase-II
FT /note= 'putative CAM kinase II phosphorylation
FT site'
FT modified_site 597
FT /label= A-kinase
FT /note= 'putative protein kinase A phosphorylation
FT site'
FT modified_site 619
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'

FT modified_site 690
FT /label= Casein_kinase-II
FT /note= 'putative casein kinase II phosphorylation
FT site'
FT modified_site 719
FT /label= CAM_kinase-II
FT /note= 'putative CAM kinase II phosphorylation
FT site'
FT modified_site 732
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 761
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 763
FT /label= A-kinase
FT /note= 'putative protein kinase A phosphorylation
FT site'
FT modified_site 765
FT /label= Casein_kinase-II
FT /note= 'putative casein kinase II phosphorylation
FT site'
FT modified_site 837
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 885
FT /label= Casein_kinase-II
FT /note= 'putative casein kinase II phosphorylation
FT site'
FT modified_site 919
FT /label= A-kinase
FT /note= 'putative protein kinase A phosphorylation
FT site'
FT modified_site 971
FT /label= CAM_kinase-II
FT /note= 'putative CAM kinase II phosphorylation
FT site'
FT modified_site 972
FT /label= A-kinase
FT /note= 'putative protein kinase A phosphorylation
FT site'
FT modified_site 972
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 1002
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 1019
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 1002
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 1105
FT /label= C-kinase
FT /note= 'putative protein kinase C phosphorylation
FT site'
FT modified_site 1194
FT /label= CAM_kinase-II
FT /note= 'putative CAM kinase II phosphorylation
FT site'
FT modified_site 1199
FT /label= Casein_kinase-II
FT /note= 'putative casein kinase II phosphorylation
FT site'
FT modified_site 1211

FT /label= C-kinase
FT /note= "putative protein kinase C phosphorylation site"
FT modified_site 1212
FT /label= CAM_kinase-II
FT /note= "putative CAM kinase II phosphorylation site"
FT modified_site 1213
FT /label= A-kinase
FT /note= "putative protein kinase A phosphorylation site"
FT modified_site 1270
FT /label= CAM_kinase-II
FT /note= "putative CAM kinase II phosphorylation site"
FT modified_site 1310
FT /label= Casein_kinase-II
FT /note= "putative casein kinase II phosphorylation site"
FT modified_site 1310
FT /label= C-kinase
FT /note= "putative protein kinase C phosphorylation site"
PM W09625502-A1.
PD 22-APR-1996.
PF 14-FEB-1996; G00312.
PR 14-FEB-1995; GB-002806.
PR 11-AUG-1995; GB-016528.
PA (MEDI-) MEDICAL RES COUNCIL.
PI Antoni F, Paterson JM;
DR WPI: 96-393403/39.
DR N-PSDB: T35209.

...
Note: remainder of annotations omitted.

Query Match 4.9%; Score 95; DB 19; Length 1353;
Best Local Similarity 26.8%; Pred. No. 2.17e+01;
Matches 15; Conservative 16; Mismatches 23; Indels 2; Gaps 2;
Db 887 efetnhvftgtsavlvvhyhcnfqlgswmrslatvgqllllllhls-lcg 941
Qy 164 WFRPALLMLPLSVYKAFVMTFVHLCSLGSWARLDARV-TGLLAKHFGPCCR 218

RESULT 5

ID R89264 standard; Protein; 772 AA.
AC R89264;
DT 07-APR-1996 (first entry)
DE Heparinase-II.
KW Heparinase-II; heparin degradation; heparan sulphate degradation;
KW Escherichia coli; polymerase chain reaction; PCR; primer; toxicity;
KW cloning; vector; ribosome binding site; haemostatic;
KW blood-clotting; antibody; affinity chromatography.
OS Flavobacterium heparinum.
FH Key Location/Qualifiers
FT peptide 1..25
FT /note= "Signal peptide"
FT peptide 85..96
FT /note= "Peptide 2B (R89267)"
FT peptide 222..233
FT /note= "Peptide 2A (R89266)"
FT peptide 500..526
FT /note= "Peptide 2C (R89268)"
PM W09534635-A1.
PD 21-DEC-1995.
PF 09-JUN-1995; U07391.
PR 10-JUN-1994; US-258639.
PA (IBEX-) IBEX TECHNOLOGIES.
PA (ZIMM-) ZIMMERMANN J.
PI Bennett C, Blain F, Gu K, Musil R, Su H, Zimmermann J;
DR WPI: 96-097381/10.
DR N-PSDB: Q59226.
PT Nucleic acids encoding Flavobacterium heparinum heparinase II and

PT III - for degrading heparin and heparan sulphate, also related host
PT cells, proteins and antibodies, useful in heparinase purification.
PS Claim 14; Fig 5; 75pp; English.
CC The sequence corresponds to Flavobacterium heparinum heparinase-II,
CC which degrades heparin and heparan sulphate. The sequence is
CC isolated from a gene library in phage lambda-DASH-II in Escherichia
CC coli by polymerase chain reaction amplification using e.g. primers
CC 2-1 to 2-4 (Q99228-Q99231), which are designed based on the
CC sequences of peptides 2A and 2B. Toxicity of the gene and natural
CC selection of the host against clones with the entire sequence has
CC been circumvented by cloning sections of the gene separately in a
CC vector with a modified ribosome binding site, which increases
CC expression levels. The heparinase-II may be used to neutralise
CC anticoagulant activity. Antibodies against the protein may be used
CC to differentiate between native and recombinant enzymes, and when
CC immobilised they may be used for heparinase purification by
CC affinity chromatography.
SQ Sequence 772 AA:

Query Match 4.8%; Score 94; DB 15; Length 772;
Best Local Similarity 26.1%; Pred. No. 2.55e+01;
Matches 23; Conservative 25; Mismatches 35; Indels 5; Gaps 3;

Db 94 dffryfngqlktrvrelmalnyntkpkvgrealtslidletatfkpgadlsrglglf 153
Qy 66 DFIEGEMKXSDVW-ADLIGLNLVWNRNA--GKGEYKIMVAALGWAT--ASLIMSRCIPLW 120
Db 154 avtgagivdywcydqlkpektrfkafv 181
Qy 121 WGANGIEFQWKYIQMSIDSLVHYVIT 148

RESULT 6

ID R57760 standard; Protein; 843 AA.
AC R57760;
DT 01-AUG-1995 (first entry)
DE Lys-aminopeptidase PepN.
KW Lys-aminopeptidase; PepN; fermented food; cheese.
OS Lactobacillus delbrueckii subsp. lactis.
PN EP-633316-A.
PD 11-JAN-1995.
PF 30-JUN-1994; 401497.
PR 01-JUL-1993; GB-013586.
PA (EEC-) EEC EURO ECONOMIC COMMUNITY.
PI Klein JR, Plapp R;
DR WPI: 95-038513/06.
DR N-PSDB: Q79913.
PT Purified Lys-aminopeptidase PepN enzyme and PepN gene - useful
PT for prep. of fermented foodstuff, esp. cheese.
PS Claim 3; Page 17-21; 41pp; English.
CC A new Lys-aminopeptidase, PepN, was isolated from L. delbrueckii
CC subsp. lactis W87 (DSM 7290) and had the sequence given in
CC R57760. The pepN gene (Q79913) was isolated from a library
CC of DSM 7290 chromosomal DNA by screening for peptidolytic activity
CC in Escherichia coli ERI562 transformants. The isolated gene
CC is used for recombinant PepN production.
SQ Sequence 843 AA:

Query Match 4.8%; Score 94; DB 13; Length 843;
Best Local Similarity 21.1%; Pred. No. 2.55e+01;
Matches 8; Conservative 15; Mismatches 14; Indels 1; Gaps 1;

Db 761 ewalektvggdnefatitvisrvfktkerydeynaf 798
Qy 129 DWKTYIQMSIDSLVHYVISAQVWMTT-RDLYLHNF 165

RESULT 7

ID W20927 standard; protein; 359 AA.
AC W20927;
DT 21-JUL-1997 (first entry)
DE H. pylori surface or membrane protein, 16ae0508orf3.
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;

RESULT 10
ID R48750 standard; Protein: 349 AA.
AC R48750;
DT 07-JUN-1996 (first entry)
DE G-protein coupled mouse glucocorticoid-induced receptor protein.
KW G-protein coupled receptor; ligand binding assay; transmembrane domain;
KW psychotic disorder; schizophrenia; dopamine; cAMP; adenosine; thrombin;
KW muscarinic acetylcholine; adrenergic; endothelin; bombesin; endocrine;
KW rhodopsin; opsin; odorant; cytomegalovirus.
OS Mus musculus.
PN M09405695-A1.
PD 17-MAR-1994.
PF 09-SEP-1993; U08528.
PR 10-SEP-1992; U542326.
PI (UTNY) UNIV NEW YORK STATE.
PA Murphy RB., Schuster DI.

DR WPI; 94-101120/12.
PT Polypeptides of G-coupled receptor proteins (GPRs) - useful for
PT Binding GPR ligands or modulating GPR binding
PS Disclosure; Page 125-126; 160pp; English.
CC Proteins R48685-R48758 represent a range of G-protein coupled receptor
CC proteins selected from cAMP, adenosine, muscarinic acetylcholine,
CC adrenergic, thrombin, endothelin, bombesin, endocrine, rhodopsin, opsin,
CC odorant, cytomegaloviral and other G-protein coupled receptors. The
CC receptor proteins were used to design polypeptides, pref. based on the
CC transmembrane domains, for use in G-protein coupled receptor ligand
CC binding assays. The polypeptide fragments retain biological activity
CC such as binding a GPR ligand or modulating GPR ligand binding to a GPR
CC (see R48759-R48758, R50569-R50807 and R89189-R89195 for examples of
CC polypeptide fragments). The polypeptide fragments can be used in
CC compositions for treating subjects suffering from a pathology related to
CC a GPR abnormality e.g. a psychotic disorder such as schizophrenia.
SQ Sequence 349 AA;

Query Match 4.6%; Score 89; DB 16; Length 349;
Best Local Similarity 20.5%; Pred. No. 5.64e+01;
Matches 18; Conservative 29; Mismatches 3; Indels 3; Gaps 3;

Db 94 alavdrhgvmlkprisltkgyivlavivmtffslphalcqklftkyedivrscl 153

Qy 125 GIEFDWYIKQNSIDNSISLVHYIVASQVMTITADLYHNFPAVLLML-SVYKAFV 183

Db 154 dppfedadl-fkwyldiatfilyllpl 180

Qy 184 ETFVHLCSLGSWRLDARA-VVTGLLAL 210

RESULT 11

ID R21412 standard; Protein; 174 AA.
AC R21412;
DT 01-APR-1992 (first entry)
DE NADH dehydrogenase 6.
KW Pneumonia; assay; AIDS; immunosuppressed.
OS Pneumocystis carinii.
PN W09119005-A.
PD 12-DEC-1991.
PF 31-MAY-1991; G00869.
PR 01-JUN-1990; GB-012196.
PA (ISIS-) ISIS INNOVATION LTD.
PI Wakefield AE, Hopkin JM, Moxon ER;
DR WPI; 92-007487/01.
DR N-PSDB; Q20065.
PT New DNA sequences which act as oligo-nucleotide primers - for
PT assaying DNA sample from respiratory secretion of a patient
PT Infected with P carinii
PS Claim 3; Fig 3; 42pp; English.
CC The amino acid sequence is that of P. carinii NADH dehydrogenase
CC 6 which was translated from DNA from plasmid pAB112. See also
CC R21409-R21413 and R20056.
SQ Sequence 174 AA;

Query Match 4.5%; Score 87; DB 3; Length 174;
Best Local Similarity 31.0%; Pred. No. 7.70e+01;
Matches 13; Conservative 14; Mismatches 1; Indels 1; Gaps 1;

Db 99 ssivltlylsiprelimekyslhfsyfswenrlngpsvveil 140

Qy 167 PAVLLMLSVYKAFVMTFVHLCSLGSWA-RLDARAVVTGL 207

RESULT 12

ID W14481 standard; Protein; 293 AA.
AC W14481;
DT 28-MAY-1997 (first entry)
DE Ranp-1.
KW ranp-1; 5'-nucleotidase consensus sequence; nervous damage; 5'-AMP;
KW treat; improved activity; organ; 5'-adenosine monophosphate.
OS Rattus rattus.
PN J09040697-A.

PD 10-FEB-1997.
PF 27-JUL-1995; 192082.
PR 27-JUL-1995; JP-192082.
PA (SHIO) SHIONOGI & CO LTD.
DR WPI; 97-175708/16.
DR N-PSDB; 762693.

PT Protein containing 5'-nucleotidase consensus sequence which
PT specifically binds AMP - is useful for treatment of nervous diseases
PS Claim 2; Page 10-11; 13pp; Japanese.
CC The sequence is that of ranp-1 which contains a 5'-nucleotidase consensus
CC sequence. The DNA expression is enhanced when there is nervous damage.
CC The protein enhances 5'-nucleotidase activity and binds specifically
CC with 5'-AMP. Nervous diseases can be treated and the activity of various
CC organs can be improved by controlling the expression of the DNA and the
CC protein.
SQ Sequence 293 AA;

Query Match 4.5%; Score 87; DB 21; Length 293;
Best Local Similarity 26.2%; Pred. No. 7.70e+01;
Matches 16; Conservative 17; Mismatches 21; Indels 7; Gaps 5;

Db 64 lfsfp-fllqfipfmkykqkdkpfeqgwkclgllfnhfflqplqctgytfe 122

Qy 3 LFFHFNCFALATFFPIITTC-SGLSE-YNAFWKCVQAGV-TYLVQL---CMLPLATF 56

Db 123 f 123

Qy.....57 F 57

RESULT 13

ID R79929 standard; Protein; 417 AA.
AC R79929;
DT 09-MAY-1996 (first entry)
DE Porcine acylglucosamine-7-epimerase mutant.
KW Porcine; acylglucosamine-2-epimerase; N-acetylmannosamine;
KW N-acetylneuraminic acid; renin-binding; enzymatic production;
KW mutant.
OS Sus scrofa.
FH Key Location/Qualifiers
FT misc_difference 23
FT /note= "wild type Met subst. with Val"
FT misc_difference 27
FT /note= "wild type Leu subst. with Met"
FT misc_difference 33
FT /note= "wild type Arg subst. with Gln"
FT misc_difference 45
FT /note= "wild type Asp subst. with Glu"
FT misc_difference 71
FT /note= "wild type Lys subst. with Thr"
FT misc_difference 72
FT /note= "wild type Leu subst. with Phe"
FT misc_difference 76
FT /note= "wild type His subst. with Arg"
FT misc_difference 77
FT /note= "wild type Arg subst. with His"
FT misc_difference 78
FT /note= "wild type Pro subst. with Ala"
FT misc_difference 79
FT /note= "wild type Glu subst. with Gln"
FT misc_difference 94
FT /note= "wild type His subst. with Tyr"
FT misc_difference 101
FT /note= "wild type Glu subst. with Gly"
FT misc_difference 120
FT /note= "wild type Ser Subst. with Thr"
FT misc_difference 137
FT /note= "wild type Val subst. with Ala"
FT misc_difference 139
FT /note= "wild type Ala subst. with Gly"
FT misc_difference 141
FT /note= "wild type Ala subst. with Val"
FT misc_difference 145

PN WO9526399-A1

```
Query Match          4.5% Score 88; DB 2; Length 550;  
Best Local Similarity 29.0%; Pred.No. 6.59e+01;  
Matches    18; Conservative   18; Mismatches 20; Indels     6; Gaps      6;
```

```
Dn  269 wyggr-ysl-wsaigslsalhwfdhfqllsqahmwqdghf-lktpleknpvillalgi 325  
       |||::|| :: | : | : | : | : | : | : | : | : | : | : | : |  
Qy  120 WGARGIEFGWMYIQMSIDSSILSVHY-VASAGVMTITRYDLHNF-RPavLLMLFLSY 177
```

Db 326 wy 327

Qy 178 -Y 178

RESULT 15

ID W48271 standard; Protein; 152 AA.

AC W48271;

DT 23-JUN-1998 (first entry)

DE Rat ninjurin 1.

KW Rat; ninjurin; cellular adhesion molecule; membrane bound; tumour;

KW nerve injury induced gene; inflammation; nervous system.

OS Rattus sp.

PN W09803650-A1.

PD 29-JAN-1998.

PF 24-JUL-1997; U12210.

PR 24-JUL-1996; US-672850.

PA (UNIK) UNIV WASHINGTON.

PI Araki T, Milbrandt J;

DR WPI; 98-120775/11.

DR N-PSDB; V20665.

PT New isolated nerve injury induced (ninjurin) gene - used to develop

PT products for treating conditions involving excessive or insufficient

PT cellular adhesion, e.g. inflammation or tumours

PS Claim 5; Fig 1B; 80pp; English.

CC The present sequence represents rat ninjurin 1 (nerve injury induced).

CC The ninjurin protein (NP) plays a role in axonal regeneration of

CC peripheral nervous system (PNS) neuronal cells after injury. The

CC products can be used for developing products for treating ninjurin

CC mediated disorders including conditions involving inappropriate (i.e.

CC excessive or insufficient) cellular adhesion. Conditions involving

CC excessive cellular adhesion which may be treated include e.g.

CC inflammatory diseases such as rheumatoid arthritis, asthma, allergy

CC conditions, adult respiratory distress syndrome, inflammatory bowel

CC diseases (e.g. Crohn's disease, ulcerative colitis and regional

CC enteritis) and ophthalmic inflammatory diseases, autoimmune diseases,

CC thrombosis or inappropriate platelet aggregation conditions,

CC arteriosclerosis, reocclusion following thrombolysis, cardiovascular

CC diseases, some forms of diabetes and neoplastic disease including

CC metastasis conditions. The products can also be used to promote

CC cellular adhesion, e.g. in nerve regeneration, wound healing or

CC prosthetic implantation. The products can also be used for detection,

CC purification, diagnosis and screening assays.

SQ Sequence 152 AA;

Query Match 4.44; Score 85; DB 29; Length 152;

Best Local Similarity 27.64; Pred. No. 1.05e+02;

Matches 21; Conservative 19; Mismatches 31; Indels 5; Gaps 5;

Db 56 llmanasqlkavvegnpefaf-fvplvlisilvlqigvyliflv-kydlnnpakha 113

Qy 111 LIMSRCIPIM-VGARGIEFDWKYIOMSDNISLVRHI-VASQVWMITRIDLYNFRPA 168

Db 114 kldflamlatglvfi 129

Qy 169 VL-LLMFLSVTKAFYM 183

Search completed: Fri Jan 22 17:59:45 1999

Job time : 84 secs.

17-1268 101 RESULT 5
ID Q33298 PRELIMINARY; PRT; 232 AA.
AC Q33298;
DT 01-NOV-1996 (TREMSLREL. 01, CREATED)
DT 01-NOV-1996 (TREMSLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMSLREL. 06, LAST ANNOTATION UPDATE)
DE ALTERNATE PETS GENE PRODUCT.
GN PETS.
OS ZEA MAYS (MAIZE).
OG CHLOROPLAST.
CC EUKARYOTA; PLANTA; EMERYOPTHTA; ANGIOSPERMAE; MONOCOTILEDONEAE;
CC CYPERALES; GRAMINEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88210525.
RA ROCK C.D., BARKAN A., TAYLOR W.C.;
RL CURR. GENET. 12:69-77(1987).
CC -1- CATALYTIC ACTIVITY: $OH(2) + 2 \text{ FERRICYTOCHROME C} = Q + 2$
CC FERROCYTOCHROME C.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).

DR EMBL: X05422; G12437; -
 DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1...
 DR PFAM: PF00033; cytochrome_b_n.
 KW CHLOROPLAST; ELECTRON TRANSPORT; RESPIRATORY CHAIN; TRANSMEMBRANE;
 KW HEME.
 SQ SEQUENCE 232 AA; 25964 MW; 88742506 CRC32;

Query Match 5.84; Score 113; DB 8; Length 232;
 Best Local Similarity 23.54; Pred. No. 8.40e-02;
 Matches 20; Conservative 27; Mismatches 32; Indels 6; Gaps 6;

Db 74 YTRPTVTFSSQYIMTEANFGWLI-R-SV-HRWSASMMVLMLVFRVYLTGGFKP 130
 Qy 132 YIQMSIDSNISLVHIVASAQ-WMITRDYLNFRPAVL-LMFLSVYKAFYMETFVHL 189

Db 131 RL-TWGTGVLAVLTASFGVTGYS 154
 Qy 190 CSLGSHWRLDARAVVTGLALKHFG 214

RESULT 6
 ID Q36615 PRELIMINARY; PRT: 232 AA.
 AC Q36615;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE APOCYTOCHROME B6.
 GN PETB.
 OS ORYZA SATIVA (RICE).
 OG CHLOROPLAST.
 OC EUKARYOTA; PLANTA; EMERTOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;
 OC CYPERALES; GRAMINEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA COTE J.C., WU N.H., WU R.;
 RL PLANT MOL. BIOL. 11:873-874(1988).
 CC -1- CATALYTIC ACTIVITY: $2\text{QH} + 2\text{Fe}^{2+} + 2\text{H}^+ \rightarrow 2\text{QH} + 2\text{Fe}^{3+} + 2\text{H}_2$
 CC FERROCYTOCHROME C.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 DR EMBL: M35955; G343210; -
 DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
 DR PFAM: PF00033; cytochrome_b_n.
 KW CHLOROPLAST; ELECTRON TRANSPORT; RESPIRATORY CHAIN; TRANSMEMBRANE;
 KW HEME.
 SQ SEQUENCE 232 AA; 26014 MW; 08640FC6 CRC32;

Query Match 5.84; Score 113; DB 8; Length 232;
 Best Local Similarity 23.54; Pred. No. 8.40e-02;
 Matches 20; Conservative 27; Mismatches 32; Indels 6; Gaps 6;

Db 74 YTRPTVTFSSQYIMTEANFGWLI-R-SV-HRWSASMMVLMLVFRVYLTGGFKP 130
 Qy 132 YIQMSIDSNISLVHIVASAQ-WMITRDYLNFRPAVL-LMFLSVYKAFYMETFVHL 189

Db 131 RL-TWGTGVLAVLTASFGVTGYS 154
 Qy 190 CSLGSHWRLDARAVVTGLALKHFG 214

RESULT 7
 ID Q32730 PRELIMINARY; PRT: 300 AA.
 AC Q32730;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE ATTA2 PROTEIN.
 GN ATTA2.
 OS AGROBACTERIUM TUMEFACIENS.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
 OC RHIZOBIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C58;
 RX MEDLINE: 96359388.
 RA MATTATSE A.G., YARNALL H.A., YOUNG N.;
 RL J. BACTERIOL. 178:5302-5308(1996).
 DR EMBL: L63540; G2340071; -
 DR PFAM: PF00528; BPD_transp.
 SQ SEQUENCE 300 AA; 32288 MW; 2B642739 CRC32;

Query Match 5.64; Score 109; DB 2; Length 300;
 Best Local Similarity 25.74; Pred. No. 2.58e-01;
 Matches 27; Conservative 28; Mismatches 41; Indels 9; Gaps 8;

Db 114 LILTGCVLPIPLSVLVSRYANVILVQRTGINQALLSTKLDEPLMITTGAVIVAMT 173
 Qy 111 LIMSRL--PLMVGARGIEFDNA--YIQMSIDSNISLVHIVASAQWMI--TRYDLYENFR 166

Db 174 HVLPLFMIMPFIAPALRAIPPEYVQAMMLGS--TRLHAFRAIIVPL 217
 Qy 167 PAVLLMFLSVYKAF--YMET--FVHLG--SLGSHWRLDARAVVTGL 207

RESULT 8
 ID Q63292 PRELIMINARY; PRT: 271 AA.
 AC Q63292;
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE NADH DEHYDROGENASE SUBUNIT 5 (FRAGMENT).
 GN NDS.
 OS BYASA ALGINOS.
 OG MITOCHONDRION
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA; PTERYGOTA;
 OC LEPIDOPTERA; UNCLASSIFIED DITRITIA; BYASA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA YAGI T., SASAKI G., TAKEBE H.;
 RL J. MOL. EVOL. 0:0-0(1998).
 DR EMBL: AB013145; D1029125; -
 KW MITOCHONDRION.
 FT NON_TER 1 1
 FT NON_TER 271 271
 SQ SEQUENCE 271 AA; 31136 MW; 1CC996F1 CRC32;

Query Match 5.54; Score 106; DB 8; Length 271;
 Best Local Similarity 34.44; Pred. No. 5.88e-01;
 Matches 21; Conservative 11; Mismatches 27; Indels 2; Gaps 2;

Db 195 YDEYITMLKSMFLLMLSLVSGSSLMQLQY-YMIFPLSLKLVYITSLMGLMGYM 253
 Qy 128 FDMKTYIQMSIDSNISLVHIVASAQWMITRDYLNFRPAVLMLFLSVYKAFYMETF 186

Db 254 V 254
 Qy 187 V 187

RESULT 9
 ID Q27297 PRELIMINARY; PRT: 212 AA.
 AC Q27297;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE CONSERVED PROTEIN.
 GN MTH1229.
 OS METHANOBACTERIUM THERMAUTOTROPHICUM.
 OC ARCHAEABACTERIA; EURYARCHAEOTA; METHANOBACTERIALES;
 OC METHANOBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DELTA H;
 RX MEDLINE: 98037514.
 RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DOUBOS J.,
 RA ALDREDGE T., BASHIRZADEH R., BLARELY D., COOK R., GILBERT K.,

Query Match 5.48; Score 105; DB 10; Length 527;

Db 129 V 129
Qv 182 V 182

Search completed: Fri Jan 22 18:02:50 1999
Job time : 94 secs.

US-09-162-597-3

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.